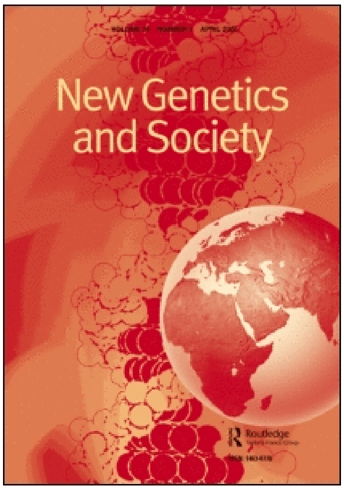

New Genetics and Society
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713439262

Genomics, ICT and the formation of R&D networks

Peter Groenewegen[a]; Paul Wouters[b]
[a] Department of Public Administration and Organization Science, Faculty of Social Science, Vrije Universiteit, 1081 HV Amsterdam, De Boelelaan 1081c, The Netherlands [b] Networked Research and Digital Information (Nerdi), The Royal Netherlands Academy of Arts and Sciences (KNAW), 1090 HC Amsterdam, The Netherlands

To cite this Article Groenewegen, Peter and Wouters, Paul(2004) 'Genomics, ICT and the formation of R&D networks', New Genetics and Society, 23: 2, 167 — 185
To link to this Article: DOI: 10.1080/1463677042000237062
URL: http://dx.doi.org/10.1080/1463677042000237062

New Genetics and Society, Vol. 23, No. 2, August 2004

Carfax Publishing
Taylor & Francis Group

Genomics, ICT and the formation of R&D networks

PETER GROENEWEGEN* & PAUL WOUTERS†
**Department of Public Administration and Organization Science, Faculty of Social Science, Vrije Universiteit, De Boelelaan 1081c, 1081 HV Amsterdam, The Netherlands, †Networked Research and Digital Information (Nerdi), The Royal Netherlands Academy of Arts and Sciences (KNAW), PO Box 95110, 1090 HC Amsterdam, The Netherlands*

ABSTRACT *Research in genomics is an example of changes induced by information and communication technologies (ICT). The emergence of interconnected ICT support for scientific work and the handling of information have changed the challenges in genomics as well as other scientific fields. The promises are significant but a large degree of uncertainty remains. While the information space is opened up, R&D cooperation essential to reaping the benefits for companies is still difficult. Moreover, in order to benefit in full from the possibility to combine knowledge on a larger scale, knowledge repositories and places of knowledge creation need to be combined. This paper discusses the new strategies of information networking between companies that emerges in response to this challenge. It concludes with an outline of a research agenda for genomics and society.*

Introduction

The use of information and communication technologies (ICT) has a broad impact on research practices in a variety of laboratory and experimental settings (National Research Council, 1997; OECD, 1999). This has pervasive effects on different levels of activity in scientific research and discovery, from discovery itself in laboratory research, the exploitation of research findings in R&D projects, to strategies in intercorporate alliances. Internet-based communication has proliferated and provided the means for engaging in new knowledge creating—interactions among researchers at the level of the research project as well as at the level of institutions and international collaborative arrangements (Buisseret & Cameron, 1994; Krück, 1995; Powell *et al.*, 1996; Zucker *et al.*, 1996). These new developments require knowledge-intensive companies, research organizations, and research facilities to have permanent access to computational capabilities and research communication networks. ICT plays a crucial role in acquiring and maintaining this access.

It is not yet entirely clear to what extent ICT have transformed actual

Correspondence to: Peter Groenewegen, Department of Public Administration and Organization Science, Faculty of Social Science, Vrije Universiteit, De Boelelaan 1081c, 1081 HV Amsterdam, The Netherlands. Tel: + 31 20 4446914. Email: p.groenewegen@fsw.vu.nl

ISSN 1463-6778 print/ISSN 1469-9915 online/04/020167-19 © 2004 Taylor & Francis Ltd
DOI: 10.1080/1463677042000237062

practices of pharmaceuticals and biotechnology companies. The role of ICT is in this field intertwined with the emergence of the increased pace of genetic research. Four years ago, Enriquez and Goldberg (2000) claimed that the experience of large companies such as Monsanto and Dupont shows that the genetic revolution has created new sector conditions. Their crucial argument is that a new science-driven dynamic has been created that forces companies to follow the flow:

> As scientific advances accelerate, more and more companies will be drawn, by choice or necessity, into the life-science business. They, too, will confront challenges unlike any they've faced before. And the way they meet those challenges will not just determine their commercial success; it will also have a direct influence over the future of life on our planet. (Enriquez & Goldberg, 2000, p. 97)

This paper appeared at the high point of optimism about the consequences of ICT applications and biotechnology in the pharmaceutical industry. Two crucial developments seemed to open up fundamentally new opportunities. First, the emergence of e-commerce carried by the internet seemed to enable new marketing strategies and business to business (B2B) communications. Second, the unravelling of the human genetic code seemed to have created a new era in which medical drugs, diagnostics, and therapeutics might be based on detailed knowledge of the underlying molecular mechanisms in the human body. Since then, however, rampant optimism has given way to more sobering thoughts. In a sense some of the technical promises mentioned above can be analysed as part of the engineering culture of promises (Van Lente & Rip, 1998). The large companies discussed by Enriquez and Goldberg have since 2000 shed some of the activities mentioned, such as their pharmaceutical divisions. They concentrate on more mundane chemical markets with their 'old fashioned' core technologies. Nevertheless, it is the pharmaceutical industry that shows the vicissitudes of the new conditions for knowledge production. The need to increase the variety of interesting products in the pipeline for pharmaceutical purposes and to garner new markets has spurred many cooperative projects in the search phase of drug discovery as well as increased investment in ICT.

The new innovative organization of knowledge-based production has been called the next phase after the Fordism-based standardization and mass customisation (Miles *et al.*, 2000). In part this reconfiguration can be developed within the existing organizations. However, in large part the formation of networks and the interconnection with public research is a necessary condition for innovative organizations (Freeman, 1991). The changes induced by international collaboration have had a marked effect on the global organisation of science and technology. In recent OECD reports, the claim is made that an important transformation of the research system is occurring. Both inward and outward research-related foreign investment has become more important. The explosion of international strategic alliances is another indicator (OECD, 1999a).

Interorganisational collaboration has interesting aspects with regards to knowledge creation, dissemination and use. These collaborations are not necessarily ICT bound. Co-operation has been described in different innovation studies on the interaction between companies and between companies and public sector research. One of the assumptions is that the industrial R&D needs direct experience and input from university and government research scientists (Faulkner *et al.*, 1995). These public-private co-operations therefore serve in part as knowledge conduits. In this paper, we concentrate on the relationships between the use of ICT in genomics research, the R&D process and pharmaceutical R&D. This field is interesting because it is a crucial area where the support of knowledge creation by ICT should have benefits (Howells, 2002).

The effects of ICT are discernible in the field of genomics at three levels: the level of laboratory practices, the level of research communication and inter-group collaboration, and the level of institutional networks. In this paper we highlight a number of the changes induced by ICT use, that might have significant impacts on the complex interactions between the research field of genetics and the commercial exploitation of science in firms, and more generally on the relationship between science and society. Our first step is a discussion of the manner in which the production and use of scientific knowledge in genomics is changing R&D. We see genomics as characterized by the merging of information technology and genetics, each with its own dynamics. We discuss how this interaction leads to a process of informatisation of research practices (Webster, 1995; Castells, 1996). Second, we discuss how the application of ICT in communication and the creation of large, distributed databases have induced a new organisation of the management of information flows in the R&D process. This development has created a new *discovery environment* in genomics, conceptualised as a socio-technical system characterised by intense interactions among human actors (scientists and technicians) and ICT (de Jong & Rip, 1997). Third, we explore how the increasing use of communication platforms is leading to the interweaving and interlocking of a variety of different hitherto separated economic actors. The field of interacting organizations itself is still in constant flux. We regard the network that is emerging as a temporal constellation. The interaction between the field level and the two underlying levels (laboratory and collaboration) is characterised by uncertainties and complex feedback processes (Orsenigo *et al.*, 2001). The paper concludes with a research agenda of these interactions that addresses the encompassing nature of the changes, possible future outcomes and some of the social issues that may arise from these developments. This research agenda focuses on the process of mediation of these institutional interactions, which has so far not yet been addressed in a systematic way in the literature on innovation. In the following sections, we will detail the aspects of this layered network of workbench, R&D integration and company networking for the case of drug development.

The informational turn in research practices

The development of genetics as a scientific and technical field has been deeply influenced by an increased role of computing and electronic communication. ICT have become crucial in various phases of the discovery process. The development of the field of bioinformatics is the direct consequence of this. In the first stage of the introduction of computing into biological sciences in the 1970s and 1980s, computer modelling and AI were introduced into biochemistry (Lenoir, 1999). Earlier on, other areas of biology such as taxonomy introduced computers because of which a gradual transformation of research work took place (Hagen, 2001). For instance, the introduction of computers as tools for repetitious tasks in taxonomy 'forced taxonomists to make their methods explicit, but just as importantly the machines served as icons of objectivity' (Hagen, 2001, p. 310). In the second stage of the widespread use of ICT, new information techniques supported the requirements of increased data management and database organisation. Both have led to the emergence of bioinformatics as a specific subdiscipline in the 1980s (Lenoir, 1999).

In what manner does ICT affect research work in the laboratory? In the chemical laboratory new substances can be 'assembled' on the basis of an integration of micro synthesis and information software tools. The emergence of combinatorial chemistry and high-through-put screening has increased the 'decision-load' of researchers. To support their decision-making capability, new tools have been developed. An example is the development of virtual compound libraries containing representations of all the molecules that might be easily obtained by the combinatorial chemistry tricks an R&D laboratory possesses (Cramer *et al.*, 1998). The use of computers has had an increasing impact on the development of research bench work in other scientific fields as well. This process has been characterised as a process of informatisation (Webster, 1995; Castells, 1996). Due to their generic aspects, the use of ICT by researchers and scholars is affecting the very core of knowledge practices across the board of all natural, technical, social and human sciences. The extent, shape and impact of these changes vary considerably, however, by scientific specialties (Kling & McKim, 2000). The generic changes in research can best be captured by the notion of the informational turn (Wouters *et al.*, 2002a; Beaulieu, 2004). First, digital information and data are playing an increasingly important role in a variety of research fields (Wouters & Schröder, 2000; Wouters *et al.*, 2002b). This has triggered a number of fields, notably the life sciences, to at least partly take on the features of data-rich information sciences (Lenoir, 1999). Second, networked research practices are becoming more abundant, although this should not be seen as diminishing the importance of local communities (Olson *et al.*, 2000; Finholt, 2002). In a number of fields, e.g., genomics, nanotechnology, and pharmaceutics, these networks link up actors from the academic as well industrial sectors of society. They are more hybrid than networks which consist solely of university-based actors.

In genetics, the developments in the first stage of computer applications have been stimulated by the sizeable investment by government agencies both in the US and Europe (Cook-Deegan, 1995). The first development relevant to our argument concerns the early characteristics of DNA sequencing research as conceived by the Human Genome Project (Jones, 2000). Technical innovations played an important role in the process. In order to complete the genetic code of human DNA large amounts of data were collected. Outcomes of sequencing needed to be stored, communicated, and combined, the data handling score became an important bottleneck. This was the basis of the first bioinformatics effort, i.e., the application of ICT to problems of biology. It also provided a first niche for scientific entrepreneurs to exploit. The support activities in *gene discovery* require access to information on sequences and function of genetic material. This has been organised by various public and commercial laboratories. Examples of universities starting with a service model for accessing gene information are Oxford, Stanford, and Sydney.

One of the interesting aspects of these new developments is that they occur in incremental steps. Each step seems to reinforce a digitisation/informatisation cycle, also leading to intervention from a higher level of organisation.

An increasing number of chemical problems can be solved through the calculation of key processes and structures, thus enabling the definition of new substances on the basis of theoretical insights (Wilson, 2000). In computational chemistry, desktop visualization of chemical activity information has expanded the possibilities of combining mathematical theoretical constructs and the chemical engineering of new substances. This requires significant computing power for interactive 3D structure visualization (Krieger, 1996). The visualisation has an important role in two different respects. First, it allows for a more intuitive use of information including structural aspects of molecules and will allow for manipulation of visualisation models. Second, there is an emergence of information based R&D integrating public information sources, organizational processes and research into a specialised information environment. This may mean that the pure information tasks are increasingly combined with the research process and organisation, through software and web based connections. Following de Jong and Rip it can be argued that the consequence of such ICT applications will be the rearrangement of various elements of the context of discovery (de Jong & Rip, 1997). This has recently been underlined by the creation of a new ‘robot scientist’ (King *et al.*, 2004) that is more powerful than the first one created in 1990 (Zytkow *et al.*, 1990). This robot is able to originate hypotheses to explain observations, devise experiments to test these hypotheses, physically run the experiments using another robot, interpret the results to falsify hypotheses inconsistent with the data, and to repeat this cycle.

An important second step in this development has been to link these data into internet- and web-based systems also encompassing links to suppliers, gene information, reagents and biological materials. Software development is crucial, in the creation of both proprietary databases and public sources, such the Human Genome Project and the SNP consortium (Thayer, 2000). On this basis, the communication work of the research process is actively integrated with the available information from different sources. The companies active in this area want to tie all relevant processes together into one system: gene discovery, target screening and selection chemical structure, and synthesis and clinical studies. The main feature is the *design of drugs*. Last, the various data are being integrated into one overall research and design environment, in which visualization tools are playing an important role. This is an indication of the networking basis that is established. A subsequent step has been the active integration of the management aspects of the research process with the available information from different sources. The service companies active in this area want to tie ERP-type systems together into one system gene discovery, target screening, selection of chemical structure and synthesis, and clinical studies. How does this emerging practice shape the conditions for corporate strategies and cooperations?

ICT, communication and collaboration in R&D: managing the new discovery environments

Much of the developments described in the previous section are an indication of the virtualisation of the organisational information infrastructure. One prominent question that arises concerns the consequences of these changes for the work processes in R&D. Thus study of communication in R&D as a social process is relevant to the distributed character of knowledge production. Allen (1977) studied information and communication patterns in corporate R&D concentrating on communication between researchers in one location. He found that members of research groups play different social roles in the communication of knowledge. Gatekeepers, not necessarily the workgroup leader, played a central role in gathering diverse information for research. They influenced what others considered reputable and relevant knowledge. Researchers who take the role of communication gatekeepers regulated access strategic information (Allen, 1977). In fulfilling this role relations, informal social connections and physical co-location played an important part in whom would fulfil this role. Also other research underlined the importance of boundary roles in the social context of a development teams as important (Tushman, 1977). The relevance of these processes is evident as the R&D process is increasingly characterized by heterogeneity of inputs as well as the internationalisation of research work within the same company (Granstrand *et al.*, 1993; Haour, 1998; Staropoli, 1998; OECD, 1999; Munari *et al.*, 2002). ICT support, exchange via expert lists, and ICT supported networked co-operation have become increasingly important:

> These [electronic and software] tools can help to get from electronics the managerial and knowledge value needed to pursue two main objectives: to use electronic space in a truly managerial way, instead of using it for the transaction of information bit; to sharpen knowledge the knowledge and information base of the firm. (Haour, 1998)

While computer supported cooperative work is very much accepted as a feature of the current work environment its impact on those organisational processes that are required for support have received scant attention from information scientists as well as social scientists.

In early studies it was argued that computer mediated communication strengthens the position of professional communities across organizational borders. This trend would increase the tension between corporations that want to reap the benefits of expertise internally and their knowledge workers (Pickering & King, 1999). The integration of new ICT tools in R&D increases the need to coordinate the R&D workflow. It raises the question: 'What tools can help to build the frames of reference needed to organize and coordinate intellectual activity?' (Bowker *et al.*, 1997, p. xvi). But the reverse is also true: in order to support cooperation, coordination and through the integration of systems, coordination of work processes at the level of companies and between organizations are a prerequisite. The ICT support processes can be distinguished in two forms: one providing the knowledge repositories for R&D processes (as discussed in the previous section), the other supporting the communication within the R&D process itself.

The improvement of communication within R&D is in all probability influenced along patterns that are similar to those found in the much wider studied work of teams in product development. Various studies found that the main enabling factor governing the adequate use of new ICT infrastructure is establishing working relationships based on personal social contacts. In absence of such trust building mechanisms, based on traditional organizational and human relations management, supporting ICT infrastructure is rather ineffective (Boutellier *et al.*, 1998). In particular the transfer of tacit knowledge is considered to be essential for the application of breakthrough information. Hansen studied cooperation in product development in an international company and found that the characteristics of the social network influenced the use of technical knowledge. Strong (social, personal friend, cooperative experience) ties in the interunit network of the company enabled the transfer tacit know how (Hansen, 1999, 2002).

Already before the emergence of ICT support mechanisms, drug development has been characterized by a significant degree of outsourcing. The need to further reduce costs is driving standardization and sub-contracting to contract research organizations. Cavalla and Gale (1997) report on a case study of a drug development project. They focused on the networking between different research, development and product organizations. The familiarity of team members with the senior project leader was found to be crucial for the integration of tasks and outcomes of subcontracted studies. The ability to personally interact at crucial moments to clarify issues and help with the interpretation of results is deemed necessary for product development. Notwithstanding this clear trend towards distributed knowledge production, it is still an open question whether successful integration can be achieved in circumstances where tasks are distributed geographically and organizationally (Cavalla & Gale, 1997). In a case study of virtual teams, Lewis concludes that management control and creating shared understandings are crucial issues. A 'flat' communication structure only goes well when participants are clear about their role (Lewis, 1998). In other studies of electronic information exchange it has been found that 'natural' structuring of the information environment occurs, with some participants taking a much more central role in the network than others (Ahuja & Carley, 1999). Thus, while in discussions of new organizational arrangements for support of virtual research environments, the suggestion of a seamless web is dominant, the outcomes of various studies suggest that emergence of directive organizing of information is essential for the realization of product goals.

To sum up, regardless of the availability of informational tools, communication networks in R&D laboratories will exhibit structural social characteristics. These structural characteristics—gatekeepers, leadership or hierarchy, affect the manner in which information gathering and communication technologies may be relevant to R&D.

Corporate innovation strategies and cooperation

The pharmaceutical industry is one of the most research-intensive sectors. Research is the cornerstone of the industry and the sector is one of the largest spenders on R&D. The development of new drugs is composed of a series of process, some of which are relatively standardized. A clear distinction is made by some authors between the research part and the development part (Chiesa, 1996). The research part is dominated by the need to search for new active chemicals. The theoretical and empirical development of pharmacology has created a situation in which a large variety of combinations of targets and chemicals can be combined in strategies for drug design (Orsenigo *et al.*, 2001). The first step is usually the identification and synthesis of chemical compounds that are considered to be biologically active. Only a small fraction of these chemical compounds completes the whole trajectory. In this planned system of discovery, innovation of the search and development process has played an important role. In this way, the pharmaceutical companies compose a field of innovation. In the 1980s and 1990s, the change towards more planned approaches was stimulated through the development of the pharmacology of receptors, as well as improvements in the laboratory instruments and automation. The number of technological fields necessary to identify chemicals for product development in the pharmaceutical industry is increasing at a rapid pace. The investment in new drug development is increasingly costly (Gambardella, 1995; Achilladelis & Antonakis, 2001). On the one hand these developments have driven a process of rapid consolidation (Heracleous & Murray, 2001). On the other hand it also led to an increase in research alliances with academia and the emergence of new research companies (Cockburn & Henderson, 1998; Orsenigo *et al.*, 2001). Orsenigo and co-authors have argued that the networking activity of the pharmaceutical industry is dependent on subsequent 'waves' of new technologies being absorbed by incumbent firms. Different waves of collaboration around drug discovery distinguish subsequent phases; for instance recombinant DNA, Monoclonal Antibodies, Rational Drug Design, Oligonucleotides, Combinatorial Chemistry have been traced. R&D cooperation of large companies focuses on new biotechnology companies, hospitals and university laboratories (Orsenigo *et al.*, 2001). New biotechnology firms have formed alliances with large pharmaceutical companies (Kenney, 1986; Grabowski & Vernon, 1994; Shan *et al.*, 1994; Powell, 1998). The emergence of high throughput screening has led towards more information and the need to select compounds within a shorter timeframe. Considerable uncertainty in the discovery process is a consequence of this change. This development was recognized by senior management but it took considerable time to overcome the internal resistance to a different manner of workbench decision making (Thomke & Kuemmerle, 2002). The identification of targets became one of the steps that were increasingly contracted out to emergent firms. In such contracts, discovery targets are set by the sponsoring pharmaceutical company that may lead to significant financial rewards for emergent firms if they reach the specified milestones. It thus seems that compared to an earlier period the social network of contracting parties itself has changed, the discovery process has become more distributed. This trend can be derived from the manner in which pharmaceutical laboratories have coped with genomics. They extended the external networking that had already started with the advent of biotechnology. In addition to the more traditional chemical and biotechnology collaborations, all kinds of information, software and information technology companies have become network cooperation parties. The broad screening programs in which many chemicals were tested on target animals or tissues, could, combined with microchips on the bench, get more refined with the help of computers in the laboratory. Examples are the labs-on-a-chip (OECD, 1999, p. 49). It is clear that the potential for digitised information exchange between various parties is increasing. Thus the pharmaceutical industry can provide a fruitful example of how genomics is regarded as an opportunity for gene-based rational drug design and the use of ICT to facilitate the exchange of relevant information in the R&D process. The development of new ICT based innovation raises three questions. First, how is the organization of the innovation process affected by these changes? Second, who is best positioned to profit from these new opportunities? Third, what might be the overall consequences of the shifts in the organization of the innovation process (Miles *et al.*, 2000)?

We consider the genomics sector as an emerging group of firms and other actors where biotechnology and information technology are being developed for innovation purposes in the agricultural, health and environmental industries. The rationalized search for new products based on genetic and additional biological and health information is increasingly based on the structured sifting of existing information and the subsequent automatic testing of potential targets and processes. The core idea of genomics as a business sector, is that increased digitisation of the discovery process and broader availability of gene and health databases offers opportunities for new businesses to enter the discovery process. In the initial phase of the evolution of genomics a lot of optimism existed about

TABLE 1. Genomic-related companies' collaboration and fate

Name	Fate	Type	Collaborations 1997–2002	Patents 1996–2001	Business Area
Incyte		A→C	11	622	Turned from information towards drug development
DoubleTwist	dissolved	A	7	0	Information provision
Celera		A→C	5	0	From data to drug development
Millennium		D	8	251	Drug development
Affymetrix		B	6	111	Sequencing hardware
CuraGen		A→C	2	20	Bioinformatics and drug discovery
Genomica	acquired	A	3	0	Software
Informax	acquired	A	2	0	Software
Tripos		A+	1	5	Software
Structural Bioinformatics		A+		1	Software
Spotfire			4	0	Planning and decision making software also outside pharmaceutical industry
Compugen		A+	3	0	Software
Lion Bioscience	reorganized	A+	3	1	Software tools and target identification

Key: (A) information based and software (B) tools for screening (C) screening and targeting

the possibility to find niches in which viable business models based on information, information structuring and retrieving functions were seen.

One consequence of this changing environment, is that new organizations and firms are stimulated to develop new business models for perceived novel opportunities. A number of ICT and biotechnology-based companies have emerged that integrate scientific, information and design software with web applications. The companies that try to profit range from information-based companies to hardware companies. In Table 1 some of the key issues are labelled for a set of companies that were identified in trade literature in 2000 as future key players (Howard, 2000; Reed, 2000). For the data-based companies, Jones describes the four basic models of information handling these companies use (Jones, 2000):

- Companies may base themselves on the possession of the tools to collect large amounts of sequencing information in a minimum of time.

- Companies that go one step further and combine the sort of data collection above with preliminary screening of biological targets.
- Other companies add insight in gene functioning to the above.
- Lastly companies emerge that relate gene information to human diseases.

Companies based on models of the exploitation of proprietary scientific and information advantages adopted a strategy to file patents based on genetic and protein information, which was made possible in the US. For instance, Human Genome Sciences has chosen a strategy initially directed at collecting the rights to large amounts of sequence data. The company Incyte provides another example in this field. It was founded in 1991 and has functioned as a provider of proprietary genome databases and analytical tools. Incyte was one of the first biotech companies to engage in high-throughput computer-aided gene sequencing for the purpose of identifying genes and their corresponding proteins with potential therapeutic applications. Incyte's database discovery approach compares partial human genes or protein sequences to genes or proteins of known sequence in order to predict their biological or therapeutic function. Prior to commencing its gene-sequencing program, Incyte used its database discovery approach to identify specific cell proteins that were expected to have pharmaceutical utility. In this way, it aimed at differentiation itself from other information providers.

Other new companies have provided commercial access with supportive software. Examples of these are Doubletwist (US, Stanford related) and eBioinformatics (US- Australia). Specialized companies also provide services such as consulting (Compugen). Through portal developments and improvements in internet transparency, researchers can access tools and search algorithms, databases, patent and scientific literature, news, and jobs listings (Thayer, 2000).

A third type of company that arose supports the R&D process through the development of tools for analysis. Examples are the producers of genechips of which Affymetrix is one example.

Lastly and of particular interest to our argument, a few companies have emerged focused at the integration of various parts of the information and planning environment of R&D. Spotfire, for instance, delivers planning and integration software that allows access to the information space and uses software to integrate and display the findings in a specific way. Thus organization of the research process such as planning, reporting and access to ongoing results has been developed as a specific set of tools.

The way in which the different levels—discovery environment, R&D collaboration, and corporate collaboration represent opportunities can be discussed on the basis of recent developments in the sector of genomic companies. When we look at the fate of the 15 companies presented in Table 1, it is clear that in the last years new patterns have emerged.

Underlying this emergence of new businesses is the need to integrate increasing amounts of data and information available on chemical, genetic and disease information. This increase has shifted the approach in pharmaceutical research from a lack of information to an overload of information on many relevant aspects. Both internal and external information sources have grown in volume and the presentation and integration of the information has become a crucial point. The surplus availability and combination of data is therefore a bottleneck in the quest for new active compounds. The approach many companies take is to shield proprietary parts of this information process. On the other hand, in order to access data sources of other companies and actors the field also has some characteristic of the field of software development. In this field an important though not uncontested premise is the strategic use of standards and open source software. In order to combine the information retrieved from databases and research these two aspects need to be combined.

It thus seems that these two fields are basic as starting elements for information centred innovation systems. In part these fields have become central, because of the necessity to integrate large amounts of diverse information that have become available. Drug research is 'data rich put information poor' (Thayer, 2000). The drug discovery process is reshaped by these changes. The manner in which these changes will affect the structure of the industry is however less clear. Fundamental differences of opinion exist between new companies as to what will constitute a viable long-term model. Initially, it was expected that data and information selling would be a viable business model in itself (Howard, 2000; Haussler, 2001). There was however also early scepticism, which claimed that pure information providers and brokers would not survive (Jones, 2000). Increasingly these functions indeed have come under pressure as the large pharmaceutical companies cut back on their spending and publicly available databases of genetic information emerged. Presently, a number of the companies in Table 1 are restructuring and some are trying to become drug developers themselves. It seems increasingly unlikely that information storing and manipulating as routine as a business model will survive. Software on decoding rearranging and interpretation still seems viable but rather marginal. Tool production as well as ERP like support for the R&D process still seems to be strong.

Some new companies, including for example US based Millennium, are hesitant to become involved in networking based on new information processes with large pharmaceutical companies and aim at building an integrated genomics based drug development capacity. Thus they may represent a new wave of emerging companies with innovation as its core and not so much centered on the new opportunities driven by. Andersen Consulting discusses the changes of large pharmaceutical companies that go in the opposite direction (Accenture, 2001). The authors claim that Web-enabled networks will further 'virtualize' R&D activities binding the partners' capabilities without regard to company borders (Accenture, 2001). The crucial bottleneck is the discovery of new drugs. The traditional screening methods fail to come up with sufficient numbers of new compounds. It is hoped that genomic information and work by other companies will help to tackle this. The use of ICT is also expected to increase access to scarce human expertise and knowledge, among others by the proliferation of internet-based clinical networks. In such networks patient groups, medical specialists and pharmaceutical companies are expected to work together. According to the Andersen report, in this process disease oriented web communities and therapy specific sites could form dual structures for testing and information exchange. The main claim of the Andersen report is that existing large companies may evolve to backbones for networks organizing information, therefore directing in the same direction (also at the high point of e-business optimism) as new emerging businesses were thought to have taken.

Mediating institutional research networks—a research agenda

Genomics has become an important field for social scientists and humanities researchers. The main focus is, however, often the ethical dimension and the exploration of possible societal consequences of the new genomics initiatives. In this section, we will argue that there is a need for studies with an additional focus. Limiting the spectrum of social studies of genomics to ethics and consequences of genomics not only runs the risk of missing important analytical dimensions of genomics and related phenomena, it also increases the danger that social scientific and humanities research will be 'frozen' in an instrumental mode, either in favour of or opposed to the further development of genomics research and applications. An effort to map the role of information and communication technologies in the development of genomics is needed, to better understand the dynamics of genomics both as a socio-technical field of *knowledge creation* and as a field of *making society*.

This approach builds further upon work in science studies and the history of science that has analysed the turn to informational research in the life sciences. Lily Kay (2000) concludes that genomics is 'moving beyond monogenetic and polygenetic determinism, even beyond functional genomics, toward a phase concerned with nonlinear, adaptive properties of complex dynamic systems, where visions of linear causality would be replaced by analyses of networks interacting with the environment and operating across levels of regulations: genetic, epigenetic, morphogenetic, and organismal'. The transformation of the life sciences into information sciences (Lenoir, 1999) makes information the essence of life. 'Genomic biopower promises new levels of control over life through the pristine metalevel of information: through control of the word, or the DNA sequence' (Kay, 2000). She therefore sees her study of the history of the book of life as a study of 'an epistemic rupture from purely material and energetic to an informational view of nature and society'. The latter is not given by nature but the product of historical contingency. It was not so much the mathematical content of information science that shaped the new agenda in the life sciences, contrary to predictions by many leaders of the informational revolution in the late 1940s, but the discourse about information. 'Information, partly displacing the concept of specificity, became a guiding metaphor, or rather a metaphor of a metaphor, in molecular biology and in the work on the genetic code' (Kay, 2000). She concludes that 'genomic textuality' has become a fact of life and commercial futures.

The informatisation has direct implications for all three levels at which genomics can be analysed. Already in 1991, Victor McKusick concluded that a new breed of scientist was being created in the genomics laboratories, 'one who is prepared to capitalize on both the molecular genetics revolution and the computation revolution. These will be the leaders in biology in the 21st century' (McKusick, 1991). The importance and steering influence of information infrastructures has been analysed by Bowker and Star (1999). Classification is not an innocent activity. The same holds for information infrastructures in general. 'Seemingly purely technical issues like how to name things and how to store data in fact constitute much of what we have come to know as natural' (Bowker & Star, 1999). This approach has been used by Mackenzie to propose two main ways to analyse the circulation of genetic sequences in bioinformatics practice (Mackenzie, 2003). One would be the analysis of the character of bioinformatics software and algorithms. The second approach would be the analysis of practices, 'the different ways that biologists, institutions and corporations retrieve and organise sequence data'. The latter would lead to an understanding of the ways in which different sets of potential determinations of living bodies are generated within an organisational milieu. This directly affects both the logic of research and the commercial and political relationships in which genomic objects are embedded and that they create. An example of the first type of implication is the invention of the improved robot scientist in the field of functional genomics (King *et al.*, 2004). As noted above, this robot is able to originate hypotheses to explain observations, devise experiments to test these hypotheses, physically run the experiments using another robot, interpret the results to falsify hypotheses inconsistent with the data, and to repeat this cycle. What Leroy Hood (2002) has coined 'discovery science' is clearly not a complete break with hypothesis driven research but extends the use of hypotheses by modelling them computationally. The potential of the robot scientist is not seen by its creators as a complete replacement of human creativity. Rather the expectation is that repetitive chores can be done by robots, thereby freeing the scientists 'to make the high-level creative leaps at which they excel'. The commercial and political aspects are important in that the new genomics databases and software tools shed light on 'how biological corporealization and the generation of new forms of property and innovation mutually contextualise themselves through movements of data, and in particular, through movements of data modulated by enterprise property relations' (Mackenzie, 2003).

Genomics research is deeply embedded in complex socio-technical configurations. ICT tools and configurations play a crucial role in the development of the field itself as well as in the interaction between genomics and society. A direct connection between the development of ICT infrastructure and the application of genomic information in research is used by the pharmaceutical industry for innovation purposes. In the previous section we introduced alternative models in which the pharmaceutical industry is rather depicted as a key organizer of health-related R&D activities, integrating genetic and other basic scientific information with health related information and health system and patient related information. In this open claim for an orchestrating role for this sector, it seems to us that also a different route is visible which starts with the need to integrate knowledge, plan research and seizing of opportunities. One direction that has been taken is the assemblage of the various knowledge elements through entrepreneurial activity. The resulting knowledge network changes the conditions for the existing corporate R&D structures. In between university science, large corporate laboratories the new start-up genomics companies emerge. In other words, the innovation processes are mediated by these networks. The role of ICT is both enabling for and constrained by these networks.

To summarize, the development of bioinformatics, enabled by internet mediation, is driving new strategies of life science research. As an example of a specific segment of players in this area, we discussed the literature with respect to the consequences of virtualisation at three levels, with the pharmaceutical industry as an example. One set of arguments concerned with these developments claims that they will result in a significant reorientation and strengthening of R&D capabilities. In line with these arguments, we can point to the emergence of companies focusing on process integration, enabling research planning and integration of available resources. Whether these companies will continue to develop as independent enterprises or as parts of big companies remains, however, an open question.

These strategies are embedded in an overall process that can be called informatisation or, more specifically, the informational turn in research (Wouters *et al.*, 2002a; Beaulieu, 2004). To grasp these developments and their impact on the niches for research strategies, it is therefore necessary to go beyond descriptive case studies of use of ICT in specific research and corporate settings. An analysis of the interaction between different levels of ICT use is needed. We have moreover concluded that rather than emerging as seamless, the networks retain specific structural characteristics. The role of *specific elements* of these networks in the dynamics of the interaction between ICT and genomics research and product development therefore becomes of crucial interest. We have characterised four different types of corporate research strategies in the literature. The effects and viabilities of these strategies as well as their effects on other users (such as patient groups) merits further study, especially with an eye to the mediating and (re)presentational role of ICT. An interesting aspect is here the balance between proprietary and public information spaces.

We propose that research should address the characterization and analysis of these dynamics at three different but interrelated levels: the level of research practice; the level of research communication and inter-group collaboration, and the level of global, institutional networks of public, academic and commercial organizations. At the first level of research practices, the workbench as it were, the role of specific elements in networks of innovation seems especially interesting. Important research questions seem to be going beyond the notion of the seamless web. In this regard, the characteristics of data flows (Arzberger *et al.*, 2004) and the balance between old and new knowledge may be interesting. At the second level of research communication and inter-group collaboration, the characteristics of inter-group networks, both in terms of structure and of process, seem especially interesting. The increased use of computer networks in communication among researchers has led to high expectations regarding 'virtual research institutes'. Such virtual institutes are presently being promoted by science policy (both at the national, European and international level) because they are expected to accelerate the process of knowledge creation. These virtual institutes are more than just the replacement of existing institutes into cyberspace. They combine functions that were previously assigned to separate social institutions, such as the library, the post office, the teleconference and the faculty. And they are built around technology that may add a new dynamic to their development. This brings to the fore questions about the influence of institutional arrangements on research practices and the way ICT affects the balance between cooperation and competition. At the third level of of global, institutional networks of public, academic and commercial organizations, the notion of niche for innovation seems especially interesting. Innovative opportunities are recognized and acted on by interdependent social actors from the private and public sector. The relations between new biotechnology firms, existing chemical and pharmaceutical companies, and universities can be analyzed as opportunity networks. Network positions influence access to resources recognition of opportunities that occur, and play a role in acquiring legitimacy (Baum *et al.*, 2000). The effects of government-led actions will be enhanced or hampered by the interaction with actions already independently undertaken by large companies, banks and venture capitalists. Therefore public resources are converted into innovation by research organisations, as well as firms (Coriat & Weinstein, 2002). Information flow and thematic focus of such networks can help to estimate the potential to exploit specific knowledge gathered/distributed in virtual research organisations. Such networks for applications in social domains need to be part of larger communities requiring balancing between public and private interests. The legitimacy of the efforts of various partners is not a priori given: it needs to be earned and secured. Relevant is the discussion on expertise on the balance between regulation aimed at risk assessment and innovation (Morris, 2000). This leads to questions about the role of specific innovation networks in genomics and the balance between the private and the public sector.

This analysis will need to draw on the sociology of knowledge, the study of the internet, and innovation network studies. The use and subsequent further development of information and communication technologies in the life sciences has consequences that reach further than the immediate instrumental aspects in the laboratory and development of data networks. They also shape the interface between genomic research and society at large. ICTs are not only configured by the user, they also 'configure the user' directly and indirectly (Woolgar, 2002). Therefore, it is crucial that the socio-technical interactions in the present take-off of genomics and bio-informatics research are analyzed. This is the more so because the use of ICTs varies by scientific field (Walsh & Bayma, 1996).

References

Accenture (2001) R&D and the Internet. Profiting from Innovative, Web-Enabled Pharmaceutical Research & Development <http://www.accenture.com/xdoc/en/industries/hls/pharma/randd_the_internet.pdf.

Achilladelis, B. & Antonakis, N. (2001) The dynamics of technological innovation: the case of the pharmaceutical industry, *Research Policy*, 30, pp. 535–88.

Ahuja, M.K. & Carley, K. (1999) Network structure in virtual organizations, *Organization Science*, 10, pp. 741–57.

Allen, T.J. (1977) *Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R&D Organization* (Cambridge, MA: MIT Press).

Arzberger, P., Schroeder, P., Beaulieu, A., Bowker, G., Casey, K., Laaksonen, L., Moorman, D., Ulhir, P. & Wouters, P. (2004) Science and government: an international framework to promote access to data, *Science*, 303(5665), pp. 1777–8.

Baum, J.A.C., Calabrese, T. & Silverman, B.S. (2000) Don't go it alone: alliance network compositions and startups' performance in Canadian biotechnology, *Strategic Management Journal*, 21, pp. 267–94.

Beaulieu, A. (Forthcoming, September 2004) From brainbank to database: the informational turn in the study of the brain, *Studies in History and Philosophy of Biological and Biomedical Sciences*.

Boutellier, R., Gassmann, O. *et al.* (1998) Managers of dispersed product development teams: the role of information technologies, *R&D Management*, 28, pp. 13–25.

Bowker, G.C. & Star, S.L. (1999) *Sorting Things Out: Classification and its Consequences* (Cambridge, MA: MIT Press).

Bowker, G.C., Turner, W., Star, S.L. & Glasser, L. (1997) Introduction, *Social Science, Technical Systems, and Cooperative Work* (Mahwah, NJ: Lawrence Erlbaum Associates).

Buisseret, T.J. & Cameron, H. (1994) Management of collaborative research: collaboration and exploitation under the UK's information engineering advanced technology programme, *Technology Analysis and Strategic Management*, 6, pp. 215–230.

Castells, M. (1996) The rise of the network society, *The Information Age: Economy, Society and Culture* (Malden, MA: Blackwell).

Cavalla, D. & Gale, D. (1997) A case history of successful virtual research, *Drug News & Perspectives: The International Drug Newsmagazine*, 10, pp. 470–5.

Chiesa, V. (1996) Separating research from development: evidence from the pharmaceutical industry, *European Management Journal*, 14, pp. 638–47.

Cockburn, I.M. & Henderson, R.M. (1998) Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery, *The Journal of Industrial Economics*, 46, pp. 157–82.

Cook-Deegan, R. (1995) *The Gene Wars. Science Politics and the Human Genome* (New York: Norton & Company).

Coriat, B. & Weinstein, O. (2002) Organizations, firms and institutions in the generation of innovation, *Research Policy*, 31, pp. 273–90.

Cramer, R.D., Patterson, D.E., Clark, R.D., Soltanstahi, F. & Lawless, M.S. (1998) Virtual compound libraries: a new approach to decision making in molecular discovery research, *Journal of Chemical Information and Computer Sciences*, 38, pp. 1010–23.

de Jong, H. & Rip, A. (1997) The computer revolution in science: steps towards the realization of computer-supported discovery environments, *Artificial Intelligence*, 91, pp. 225–6.

Enriquez, J. & Goldberg, R.A. (2000) Transforming life, transforming business: the life- science revolution, *Harvard Business Review*, pp. 96–104.

Faulkner, W., Senker, J. *et al.* (1995) *Knowledge Frontiers. Public Sector Research and Industrial Innovation in Biotechnology, Engineering Ceramics, and Parallel Computing* (Oxford: Clarendon Press).

Finholt, T.A. (2002) Collaboratories. In Shaw, D. (ed) *Annual Review of Information Science and Technology*, p. 36 (Medford, NJ: Information Today).

Freeman, C. (1991) Networks of innovators: a synthesis of research issues, *Research Policy*, 20, pp. 499–514.

Gambardella, A. (1995) *Science and Innovation. The US Pharmaceutical Industry During the 1980s* (Cambridge: Cambridge University Press).

Grabowski, H. & Vernon, J. (1994) Innovation and structural change in pharmaceuticals and biotechnology, *Industrial and Corporate Change*, 3, pp. 435–49.

Granstrand, O., Hakanson, L. & Sjolander, S. (1993) Internalization of R&D—a survey of some recent research, *Research Policy*, 22, pp. 413–30.

Hagen, J.B. (2001) The introduction of computers into systematic research in the United States during the 1960s, *Studies in the History and Philosophy of Biology and Biomedical Science*, 32, pp. 291–314.

Hansen, M.T. (1999) The search-transfer problem: the role of weak ties in sharing knowledge across organization subunits, *Administrative Science Quarterly*, 44, pp. 82–111.

Hansen, M.T. (2002) Knowledge networks: explaining effective knowledge sharing in multiunit companies, *Organization Science*, 13, pp. 232–48.

Haour, G. (1998) The global village of innovation and R&D management—an introduction, *R&D Management*, 28, pp. 1–4.

Haussler, D. (2001) The challenge of bioinformatics, *R&D Magazine*, 43, 8S–SC3.

Heracleous, L. & Murray, J. (2001) The urge to merge in the pharmaceutical industry, *European Management Journal*, 19, pp. 430–7.

Hood, L. (2002) *My Life and Adventures Integrating Biology and Technology*. A Commemorative Lecture for the 2002 Kyoto Prize in Advanced Technologies <http://www.systemsbiology.org/extra/2002Kyoto.pdf.

Howard, K. (2000) The bioinformatics gold rush, *Scientific American*, July, pp. 58–63.

Howells, J. (2002) Mind the gap: information and communication technologies, knowledge activities and innovation in the pharmaceutical industry, *Technology Analyis and Strategic Management*, 14, pp. 355–70l.

Jones, P.B.C. (2000) The commercialization of bioinformatics, *Electronic Journal of Biotechnology*, 3, p. 20.

Kay, L.E. (2000) *Who Wrote the Book of Life? A History of the Genetic Code* (Stanford: Stanford University Press).

Kenney, M. (1986) Schumpetrian innovation and entrepreneurs in capitalism: a case study of the US biotechnology industry, *Research Policy*, 15, pp. 21–31.

King, R.D., Whelan, K.E., Jones, F.M., Reiser, P.G., Bryant, C.H., Muggleton, S.H., Kell, D.B. & Oliver, S.G. (2004) Functional genomic hypothesis generation and experimentation by a robot scientist, *Nature*, 427, pp. 247–52.

Kling, R. & McKim, G. (2000) Not just a matter of time: field differences in the shaping of electronic media in supporting scientific communication, *Journal of the American Society for Information Science*, 51, pp. 1306–20.

Krieger, J.H. (1996) Genomic information, web technology drive chemical software development, *Chemical and Engineering News*, 17, pp. 10–17.

Krück, C.P. (1995) EUREKA: the construction of international R&D cooperation. In Thorn, Christopher (ed) *Benevolent Conspiracies. The role of Enabling Technologies in the Welfare of Nations. The Cases of SDI, SEMATECH and EUREKA*, pp. 175–324. (Berlin: Walter de Gruyter).

Lenoir, T. (1999) Shaping biomedicine as an information science. Proceedings of the 1998 on the History and Heritage of Science Information Systems. In Bowden, M.E., Bellardo Hahn, T. & William, R.V. (eds) *ASIS Monograph Series* (Medford, NJ: Information Today Inc.).

Lewis, R. (1998) Membership and management of a 'virtual' team: the perspectives of a research manager, *R&D Management*, 28, pp. 5–12.

Mackenzie, A. (2003) Bringing sequences to life, *New Genetics and Society*, 22(3), pp. 315–8.

McKusick, V. (1991) Current trends in mapping human genes, *FASEB Journal*, 5, p. 19.

Miles, R.E., Snow, C.C. & Miles, G. (2000) The future.org., *Long Range Planning*, 33, pp. 300–21.

Morris, N. (2000) Vial bodies: conflicting interests in the move to new institutionalrelationships in biological medicines research and regulation, *Research Policy*, 29(2), pp. 149–67.

Munari, F., Roberts, E.B. & Sobrero, M. (2002) Privatization processes and the redefinition of corporate R&D boundaries, *Research Policy*, 31(1), pp. 31–53.

National Research Council (1997) *Bits of Power. Issues in Access to Scientific Data* (Washington, DC: National Academy Press).

OECD (1999a) *Global Research Village* (Paris: OECD).

OECD (1999b) *Globalisation of Industrial R&D: Policy Issues* (Paris: OECD).

Olson, G., Finholt, T. *et al.* (2000) *Behavioral Aspects of Collaboratories*, pp. 1–14 (Mahwah, NJ: Lawrence Erlbaum Associates).

Orsenigo, L., Pammoli, F. & Massimo, R. (2001) Technological change and network dynamics: lessons from the pharmaceutical industry, *Research Policy*, 30(3), pp. 485–508.

Pickering, J.M. & King, J.L. (1999) Hardwiring weak ties. Interorganizational computer-mediated communication, occupational communities, and organizational change. In Fulk, J. (ed) *Shaping Organization Form. Communication, Connection, and Community*, pp. 399–413 (Thousand Oaks: Sage).

Powell, W.W. (1998) Learning from collaboration: knowledge and networks in the biotechnology and pharmaceutical industries, *California Management Review*, 40, pp. 228–40.

Powell, W.W., Koput, K.W. & Smith-Doerr, L. (1996) Interorganizational collaboration and the locus of innovation: networks of learning in biotechnology, *Administrative Science Quarterly*, 41, pp. 116–45.

Reed, J. (2000) *Trends in Commercial Biotechnology* (New York: Oscar Gruss).

Shan, W., Walker, G. & Kogut, B. (1994) Interfirm cooperation and startup innovation in the biotechnology industry, *Strategic Management Journal*, 15, pp. 387–94.

Staropoli, C. (1998) Cooperation in R&D in the pharmaceutical industry—the network as an organizational innovation governing technological innovation, *Technovation*, 18, pp. 13–23.

Thayer, A.M. (2000) Bioinformatics for the masses, *Chemical and Engineering News*, 78, pp. 19–32.

Thomke, S. & Kuemmerle, W. (2002) Asset accumulation, interdependence and technological change: evidence form pharmaceutical drug discovery, *Strategic Management Journal*, 23(7), pp. 619–35.

Tushman, M.L. (1977) Special boundary roles in the innovation process, *Administrative Science Quarterly*, 22, pp. 587–604.

Van Lente, H. & Rip, A. (1998) The rise of membrane technology. From rhetorics to social reality, *Social Studies of Science*, 28, pp. 221–54.

Walsh, J.P. & Bayma, T. (1996) Computer networks and scientific work, *Social Studies of Science*, 26(3), pp. 661–703.

Webster, F. (1995) *Theories of the Information Society* (London: Routledge Kegan Paul).

Wilson, E.K. (2000) Chemists turned visionaries, *Chemical and Engineering News*, 78, pp. 39–45.

Woolgar, S. (2002) *Virtual Society? Technology, Cyberbole, Reality* (Oxford: Oxford University Press).

Wouters, P., Beaulieu, A. & Woo Park, H. (2002a) Knowledge production in the new digital networks. Network Research and Digital Information (Nerdi) NIWI-KNAW <http://www.niwi.knaw.nl/en/nerdi2/research_programme/toon

Wouters, P., Beaulieu, A. & Reddy, C. (2002b) Networks and databases: tools for the greater good? 26th Annual Meeting of 4S. Milwaukee, USA.

Wouters, P. & Schröder, P. (2000) Access to publicly financed research: the global research village III. [Amsterdam: Organisation for Economic Co-operation and Development, Netherlands Ministry of Education, Culture and Science (Min. OCW). NIWI-KNAW].

Zucker, L.G., Darby, M.R., Brewer, M.B. & Peng, Y. (1996) Collaboration structure and information dilemmas in biotechnology. In Tyler, T.R. (ed) *Trust in Organizations. Frontiers of Theory and Research*, pp. 90–113 (Thousand Oaks: Sage).

Zytkow, J.M., Zhu, J. & Hussam, A. (1990) *Automated Discovery in a Chemistry Laboratory*. 8th National Conference on Artificial Intelligence (AAAI–1990) (Cambridge, MA: MIT).